



POLITECNICO  
DI TORINO

SmartData@PoliTO



Thursday, April 11<sup>th</sup>, 2019 16:30

Aula AI DISAT

Corso Duca degli Abruzzi, 24

# Francesco Bonchi

ISI Foundation, Turin, Italy



## *Distance-based community search*

### ABSTRACT

Suppose we have identified a set of subjects in a terrorist network suspected of organizing an attack. Which other subjects, likely to be involved, should we keep under control? Similarly, given a set of patients infected with a viral disease, which other people should we monitor? Given a set of proteins of interest, which other proteins participate in pathways with them? Each of these questions can be modeled as a graph-query problem: given a graph  $G = (V, E)$  and a set of query vertices  $Q$ , find a subgraph  $H$  of  $G$  which “explains” the connections existing among the nodes in  $Q$ , that is to say that  $H$  must be connected and contain all query vertices in  $Q$ .

We start by providing a brief survey of various measures and methods defined for this network problem, then we turn our attention to the problem of finding a “minimum Wiener connector”, i.e., the subgraph of  $G$  that connects all query vertices and that minimizes the sum of all pairwise shortest-path distances between its vertices (Wiener Index). We show that the minimum Wiener connector is smaller and denser than other methods in the literature, and it contains highly central nodes.

In the second part of the talk, we relax the constraint of connecting all the query vertices. Relaxing the connectedness requirement allows the connector to detect multiple communities and to be tolerant to outliers. We achieve this by introducing the new measure of network inefficiency and by instantiating our search for a selective connector as the problem of finding the minimum inefficiency subgraph. We show that our problem is hard and devise efficient algorithms to approximate it. By means of several case studies in a variety of application domains (such as human brain, cancer, and food networks), we show that our minimum inefficiency subgraph produces high-quality solutions, exhibiting all the desired behaviors of a selective connector.

Finally, we extend the present notions to the case of temporal dynamic networks showing how our tools can be used to track a community of interest adaptively in time.

### BIOGRAPHY

Francesco Bonchi is Deputy Director at the ISI Foundation, Turin, Italy, with responsibility over the Industrial Research area. At ISI Foundation, he is also Research Leader for the “Algorithmic Data Analytics” group. He is also (part-time) Research Director for Big Data & Data Science at Eurecat (Technological Center of Catalunya), Barcelona. He was Director of Research at Yahoo Labs in Barcelona, Spain. He has more than 200 publications in these areas. He also filed 15 US patents, and got granted 8 US patents. More info at <http://www.francescobonchi.com/>

<https://smartdata.polito.it/category/seminars/>

